¿Conectan los futuros profesores las aproximaciones frecuencial y clásica de la probabilidad?

Rafael Parraguez[1], María M. Gea[2], Danilo Díaz-Levicoy[3] y Carmen Batanero[4]

Resumen

En la última década, el currículo de muchos países ha incorporado contenidos de probabilidad desde los primeros niveles de educación primaria, lo que puede suponer un reto para el profesorado encargado de su enseñanza. Los programas educativos sugieren utilizar un enfoque frecuencial de la probabilidad, donde se realicen experimentos, recojan datos de sus resultados y obtengan conclusiones sobre la probabilidad de diferentes sucesos. En este trabajo se analiza la forma en que los futuros profesores de educación primaria en España relacionan los significados clásicos y frecuencial de la probabilidad y las dificultades que puedan tener en diferenciar y poner en relación dichos significados. Entre los resultados obtenidos destacamos, la dificultad en estimar la frecuencia esperada de veces que ocurre un suceso.

Palabras clave: Probabilidad (significado frecuencial y clásico), Futuros profesores.

Abstract

In the last decade, the curriculum of many countries has incorporated probabilistic content from the first levels of primary school, which can be challenging for teachers in charge of teaching. Educational programs suggest using a frequentist approach to probability, where children conduct experiments, collect data and obtain conclusions on the likelihood of different events. In this work the way prospective primary school teachers in Spain relate classical and frequency meanings of probability, as well as the difficulties they have in differentiating these meanings and relate them are analyzed. Among the results we highlight the difficulty in estimating the expected frequency of times an event occurs.

Keywords: Probability (Frequency and Classic Meaning), Prospective teachers.

I. Introducción

Un cambio importante que se ha producido en la última década en el currículo de educación primaria de muchos países es la incorporación de contenidos de probabilidad (e.g., MECD, 2014). La principal razón que apoya este cambio es que vivimos en un

[1] Universidad de Granada, España. rafparra@correo.ugr.es
[2] Universidad de Granada, España. mmgea@ugr.es
[3] Universidad de Granada, España. dddiaz01@hotmail.com
[4] Universidad de Granada, España. batanero@ugr.es

mundo con fuerte presencia del azar, por lo que hemos de preparar a los niños y niñas para afrontar dichas situaciones y tomar decisiones correctas. Así, autores como Gal (2005) reclaman la alfabetización probabilística de todos los ciudadanos, entendida como el conjunto de conocimientos, capacidades y actitudes que les permitan desenvolverse ante fenómenos aleatorios. Será entonces necesario proporcionar una formación adecuada de los futuros profesores que han de impartir estos contenidos, teniendo en cuenta las características específicas de la probabilidad y sus aspectos didácticos.

En este trabajo nos hemos interesado por esta problemática y más concretamente, por analizar la forma en que los futuros profesores de este nivel educativo relacionan los significados clásicos y frecuencial de la probabilidad y las dificultades que puedan tener al diferenciar y poner en relación dichos significados. Para ello se analizan las respuestas escritas a una actividad práctica resuelta por una muestra de futuros profesores de educación primaria en España. Dicha actividad se ha tomado de un trabajo previo de Rivas y Godino (2015), quienes también la utilizaron en la formación de profesores. Nuestro análisis completa el suyo, puesto que los autores citados sólo describen la resolución colectiva en la pizarra por el profesor y alumnos. En nuestro caso, comparamos las respuestas en un grupo de estudiantes que la resuelven individualmente y otro que la resuelven trabajando en parejas, en ambos casos por escrito.

II. Fundamentos

El concepto de probabilidad tiene diversas interpretaciones, aunque en el presente trabajo nos centramos en las definiciones de probabilidad clásica y frecuencial, por su papel en el currículo de educación primaria. Según Batanero (2005), ambos han de ser comprendidos por los estudiantes y, en nuestro caso, los futuros profesores, quienes deben ponerlos en relación para que puedan ser enseñados de manera significativa al alumnado.

La interpretación clásica tiene su origen en los juegos de azar; la correspondencia entre Pascal y Fermat en la década de 1650 se considera el primer estudio matemático de la probabilidad, aunque en esta correspondencia todavía no se presenta una definición del concepto. La primera definición formal la entrega de Moivre (1967) en *The Doctrine of Chances*:

> Si constituimos una fracción cuyo numerador es el número de chances (posibilidades) con la que el suceso podría ocurrir y el denominador el número de chances con las que puede ocurrir o fallar, esta fracción será una definición propia de la probabilidad de ocurrencia (p. 1).

Seguido a esto, Laplace (1995) publica un texto donde estableció la definición que actualmente conocemos como probabilidad clásica o regla de Laplace: “como una fracción cuyo numerador es el número de casos favorables y cuyo denominador el número de todos los casos posibles” (p. 28). Esta definición, desde su comienzo, no estuvo ajena a la controversia.

Según Godino, Batanero y Cañizares (1987), la definición es circular, pues el término “equiprobable” se incluye en la definición de probabilidad; además, solo se puede aplicar esta definición a experimentos con un número finito de posibilidades. Esta definición se usa en la escuela, porque a los niños les gustan los juegos y de este modo se interesan por el tema, pero el rango de aplicaciones que se puede mostrar con este enfoque es muy limitado.

Al tratar de superar los inconvenientes y controversias filosóficas del enfoque clásico, algunos autores definen la probabilidad desde un punto de vista frecuencial. Los estudios sobre tablas de vida en Inglaterra, en que se recopilaron muchos datos sobre edad y causas de muerte, habían hecho observar que la frecuencia relativa de un suceso (por ejemplo, una causa de muerte) se estabiliza en el tiempo (Batanero, Henry y Parzysz, 2005). El primero en demostrar que la probabilidad de un suceso se puede estimar con la precisión que se desee a partir de la frecuencia relativa, observada en una serie grande de ensayos del mismo experimento, fue Bernoulli (1987). La ley débil de los grandes números establece que dados $\varepsilon > 0$ y $\alpha > 0$ tan pequeños como queramos, podemos encontrar un valor $n > \frac{pq}{\varepsilon\alpha^2}$, siendo p la probabilidad teórica y $q = 1 - p$, de forma que se verifica que la diferencia entre la frecuencia relativa h_n en n ensayos y la probabilidad cumple que: $P(|h_n - p| < \varepsilon) \geq 1 - \alpha$.

La convergencia de la frecuencia relativa a la probabilidad teórica se consideró como un apoyo al carácter objetivo de la probabilidad, pero hubo que esperar a Von Mises (1952), quien fue el primero en definir la probabilidad en su concepción frecuencial como el valor hacia el cual tiende la frecuencia relativa en un gran número de ensayos.

Esta definición amplía la posibilidad de aplicar el cálculo de probabilidades, pues no se exige la equiprobabilidad de los sucesos elementales. Con ello, la estadística y probabilidad comienzan a aplicarse en todos los campos de las ciencias y ramas de la actividad humana (Godino, Batanero y Cañizares, 1987). Sin embargo, esta definición presenta algunos problemas. El primero de ellos es que, al partir de la frecuencia relativa, no obtenemos un valor exacto de la probabilidad, sino una estimación de la misma, por tanto, puede variar cada vez que tratamos de estimarla. La segunda cuestión es que, en la mayor parte de los casos, es imposible realizar los experimentos exactamente en las mismas condiciones y hay circunstancias (por ejemplo, en medicina, en inversiones, etc.) donde con seguridad sabemos que las condiciones de una y otra repetición son diferentes. Por último, es difícil comprender cuál es el número de experimentos que debemos realizar para que el valor de la probabilidad estimado sea válido o considerado bueno.

Según Godino, Batanero y Cañizares (1987), en la enseñanza, esta definición permite establecer la unión entre la estadística y la probabilidad, pues utiliza el concepto de frecuencia relativa de la estadística aplicándolo en el cálculo de probabilidades. Además, gracias a la tecnología, es muy sencillo aplicar este enfoque, usando la simulación para repetir un cierto experimento un número grande de veces y observar empíricamente la convergencia. Se aborda así la dificultad de la ley de los grandes números, sustituyéndola por una aproximación empírica e intuitiva de la misma.

El futuro profesor, que como recomienda el currículo ha de enseñar en la educación primaria con estos dos enfoques, ha de diferenciarlos claramente y dominar los elementos que caracterizan a cada uno. También debe relacionarlos, comprendiendo las exigencias de cada enfoque, la diferencia entre estimación a partir de la frecuencia relativa y probabilidad teórica. Esta comprensión relacional es la que tratamos de evaluar en nuestro trabajo.

III. Antecedentes

Las investigaciones previas sobre los conocimientos de la probabilidad de los futuros profesores de educación primaria indican dificultades en el tema, lo cual es hasta cierto punto razonable, ya que la probabilidad sólo se ha incluido recientemente en el currículo escolar desde los primeros niveles. Por ello, no se consideraba necesario formar a los futuros profesores de educación primaria sobre este tema en las Escuelas de Formación del Profesorado. A continuación, destacamos las más relacionadas con nuestro estudio.

Azcárate (1995), que estudió las respuestas a un cuestionario de 57 futuros profesores de educación primaria, muestra que la mayoría relacionaron la aleatoriedad con la causalidad, no comprendiendo la utilidad de la probabilidad para estudiar los fenómenos aleatorios. También observó el predomino de esquemas causales, así como dificultades en el uso y comprensión de información frecuencial en la cuantificación de probabilidades y en la idea de juego equitativo.

Serrano (1996) plantea, a una muestra de 130 futuros profesores, un cuestionario para evaluar tres componentes de su conocimiento sobre la probabilidad: a) las propiedades atribuidas a las secuencias de resultados aleatorios; b) la interpretación de enunciados de probabilidad frecuencial; y c) el uso de heurísticas en la resolución de problemas probabilísticos sencillos. Encontró una fuerte presencia de la heurística de la representatividad, es decir, esperar la convergencia en muestras pequeñas (Kahneman, Slovic y Tversky, 1982) y del enfoque en el resultado (Konold, 1989), es decir, interpretar un problema de probabilidad en forma no probabilística.

Batanero, Cañizares y Godino (2005) estudiaron la influencia de actividades de simulación sobre el conocimiento común de la probabilidad en 132 futuros profesores españoles de educación primaria. En las respuestas de los participantes al cuestionario propuesto, los autores observaron la heurística de la representatividad; el sesgo de equiprobabilidad (Lecoutre, 1992), consistente en suponer todos los sucesos aleatorios como equiprobables; y el enfoque en el resultado (23% de estudiantes). Estos sesgos disminuyen bastante después de unas actividades de simulación, en que los estudiantes interaccionan directamente con el problema gracias al uso de la tecnología.

Mohamed (2012) aplicó un cuestionario, usado previamente con niños, a 283 futuros profesores de educación primaria donde las preguntas más sencillas fueron asignación de probabilidades simples, probabilidad condicionada, estimación frecuencial de la probabilidad en el corto plazo y juego equitativo en un experimento simple. Los principales errores fueron la confusión entre suceso seguro y posible, la falta de razonamiento combinatorio o ideas confusas sobre características de los fenómenos aleatorios como la falta de percepción de la independencia y dificultades asociadas al experimento compuesto.

El trabajo más completo sobre el tema fue realizado por Gómez (2014). Con base a un estudio detallado del currículo y los libros de texto, así como de los diferentes significados de la probabilidad, construye un cuestionario con una metodología rigurosa. La aplicación del cuestionario en una muestra de 157 futuros profesores y el análisis detallado de sus respuestas permite caracterizar dichos componentes de su conocimiento. Además, evalúa el conocimiento diferenciado de las aproximaciones clásica, frecuencial y subjetiva de la probabilidad. Por otro lado, las dificultades encontradas en la literatura sobre el tema son retomadas en una actividad formativa, basada en el debate de las soluciones y la simulación para desarrollar el conocimiento matemático de la probabilidad en estos futuros profesores. En los resultados obtenidos, los futuros profesores presentaron un conocimiento común adecuado respecto al enfoque clásico de la probabilidad; por ejemplo, al enumerar el espacio muestral y en el cálculo de probabilidades sencillas. Presentaron mayores dificultades en el desempeño del enfoque frecuencial, donde muchos participantes no perciben la variabilidad de las pequeñas muestras o caen en el sesgo de equiprobabilidad.

Aunque no hemos encontrado investigaciones sobre futuros profesores de educación primaria en que se analice cómo integran los significados clásico y frecuencial de la probabilidad, sí hemos encontrado algunas investigaciones con estudiantes. Entre ellas destacamos la de Valdés (2016), en que analiza el trabajo de 30 estudiantes cuando abordan una situación de muestreo en que deben predecir o bien la composición de la población, o bien los resultados en diferentes muestras. El autor caracteriza las ideas informales de aleatoriedad, variabilidad e independencia en estos estudiantes.

Para completar las investigaciones previas, en este trabajo nos centramos en una muestra de futuros profesores que, mediante un cuestionario escrito, se enfrentan a un conflicto en el que se les ofrecen datos empíricos sobre un juego, que contradicen las probabilidades calculadas en sentido clásico. Esta situación motivará la reflexión entre las aproximaciones clásica y frecuencial de la probabilidad, que es uno de nuestros propósitos.

IV. Método

La muestra estuvo compuesta por 60 estudiantes, de los cuales 20 realizaron la actividad en forma individual y los 40 restantes en parejas (un total de 20 parejas). Estos estudiantes cursaban el segundo año del Grado de Maestro de educación primaria en la Universidad de Granada y en el año anterior cursaron la asignatura *Bases Matemáticas para la Educación Primaria*, en que se estudian los contenidos relativos a la probabilidad durante aproximadamente una semana. La recogida de datos formó parte de una práctica realizada en el transcurso de la asignatura de segundo curso denominada *Enseñanza y Aprendizaje de las Matemáticas en Educación Primaria*, cuando ya habían estudiado la mayoría del temario centrado en la Didáctica de la Matemática.

Para explicar la situación y facilitar la recogida de datos, se dio a los estudiantes una ficha de trabajo como se muestra en la Figura 1, en la que se describe una secuencia de enseñanza de la probabilidad llevada a cabo por una maestra, que es tomada de la investigación de Rivas y Godino (2015). Los estudiantes debieron responder por escrito a cada pregunta de la actividad y, además de la situación inicial presentada en la Figura 1, también contestar a las preguntas siguientes:

1. Analiza la actividad 1. Indica quién tiene ventaja en el juego y por qué. Determina la probabilidad teórica.
2. Si realizáramos muchas tiradas, ¿Con cuánta frecuencia esperarías que saliera cada una de las sumas? ¿Con cuánta frecuencia ganaría *A* o *B*?
3. Analiza la Actividad 2 ¿Piensas que la tabla con los datos de toda la clase es suficiente para sacar conclusiones del experimento realizado y las cuestiones planteadas? ¿Podríamos tomar como probabilidad de ganar *A* la frecuencia relativa de veces que gana *A*?

Raquel es maestra de Primaria y lleva unos días trabajando algunas nociones de probabilidad en su clase de 6º. Sus estudiantes asignan sin dificultad probabilidades a ciertos sucesos simples, como los resultados de lanzar un dado o una moneda o de girar una ruleta con todos los sectores iguales. Para hoy decide plantear la siguiente actividad:

Actividad 1. Suma de puntos al lanzar dos dados:

"*Vamos a jugar con dos dados por parejas. Lanzamos los dados y sumamos los puntos obtenidos. Si resulta una suma de 6, 7, 8, o 9 entonces gana A una ficha; si la suma es distinta de esos números gana B una ficha. ¿Qué prefieres ser jugador A o B?*

Juega con un compañero 10 veces y anota los resultados de las sumas que obtienes. ¿Quién ha ganado más veces A o B? ¿Piensas que se repetirá el resultado si jugamos 100 veces más? ¿Por qué?

Actividad 2. Recogida de datos de la clase:

A continuación Raquel recoge en la siguiente tabla los datos de las 10 parejas de estudiantes que hay en la clase y les pide construir un diagrama de barras con estos datos.

	Suma de puntos	Número de veces	Frecuencia relativa
Gana B	2	2	0,02
	3	9	0,09
	4	12	0,12
	5	20	0,2
Gana A	6	7	0,07
	7	12	0,12
	8	14	0,14
	9	9	0,09
Gana B	10	8	0,08
	11	4	0,04
	12	3	0,03

Raquel plantea a los niños las siguientes preguntas: ¿Quién ha ganado más veces los jugadores *A* o los *B*? ¿Quién tiene más probabilidad de ganar?

Figura 1. Ficha de trabajo de los estudiantes

Nuestro análisis se centra en las respuestas de los futuros profesores a estas cuestiones, ya que nos interesamos por la comprensión relacional de estos estudiantes en las aproximaciones clásica y frecuencial de la probabilidad. Previamente, y para elaborar sus argumentaciones, los futuros profesores deberán haber resuelto las dos actividades que componen la situación de enseñanza planteada (Figura 1).

V. Resultados y discusión

Análisis del juego desde el significado clásico. Aproximación intuitiva al enfoque frecuencial

1. Analiza la actividad 1. Indica quién tiene ventaja en el juego y por qué. Determina la probabilidad teórica.

Para llegar a la solución de la actividad 1 (Figura 1), el futuro profesor podría utilizar algún tipo de esquema basado en la enumeración sistemática del espacio muestral del experimento, como el mostrado en la Tabla 1. Se observa que no todas las sumas de puntos de los dados tienen la misma probabilidad de ocurrir, siendo más probables las sumas intermedias, que justamente corresponden a las posibilidades de ganar el jugador *A*; siendo menos probables las posibilidades extremas. Puesto que, aplicando el principio de indiferencia, todos los sucesos elementales tienen igual probabilidad de ocurrir, a partir del esquema dado se deduce que *B* gana 16 veces de 36. Con un razonamiento semejante se deduce que *A* gana 20 veces de 36. Utilizando la regla de Laplace, como cociente entre casos favorables y posibles, se deducen las probabilidades teóricas de A: P(*A*)= 20/36=0,55 y de B: P(*B*)= 16/36= 0,45. Por tanto, *A* tiene ventaja en este juego. Se espera que los futuros profesores posean una comprensión, al menos intuitiva, de la ley de los grandes números cuando argumenten que la variabilidad de éxito en la ganancia en 10 jugadas no se tiene que repetir si el experimento se realizase unas 100 veces.

Tabla 1. *Enumeración sistemática para resolver la tarea 1*

	Suma de puntos	Casos: (dado1, dado2)
Gana B	2	(1,1)
	3	(1,2); (2,1)
	4	(1,3); (2,2); (3,1)
	5	(1,4); (2,3); (3,2); (4,1)
Gana A	6	(1,5); (2,4); (3,3); (4,2); (5,1)
	7	(1,6); (2,5); (3,4); (4,3); (5,2); (6,1)
	8	(2,6); (3,5); (4,4); (5,3); (6,2)
	9	(3,6); (4,5); (5,4); (6,3)
Gana B	10	(4,6); (5,5); (6,4)
	11	(5,6); (6,5)
	12	(6,6)

Hemos encontrado las siguientes categorías de respuestas:

R1. Respuesta correcta. Se identifican correctamente todos los sucesos del espacio muestral producto en el lanzamiento de dos dados, haciendo una enumeración sistemática y teniendo en cuenta el orden. A partir de ello, se identifican correctamente los sucesos correspondientes a las sumas en las que ganan *A* y *B*, y se calcula las probabilidades correspondientes aplicando la regla de la suma, o bien el sucesos complementario. Mostramos a continuación un ejemplo (Figura 2), en que el estudiante usa una tabla de doble entrada para formar el espacio muestral producto. En el cuerpo de la tabla, introduce una notación numérica para identificar cuáles de ellos tienen la misma suma. Con dos colores marca los sucesos favorables al jugador *A* o *B*, según corresponda, que expresa mediante un código situado al margen de la tabla. Además, calcula las probabilidades aplicando la regla de Laplace.

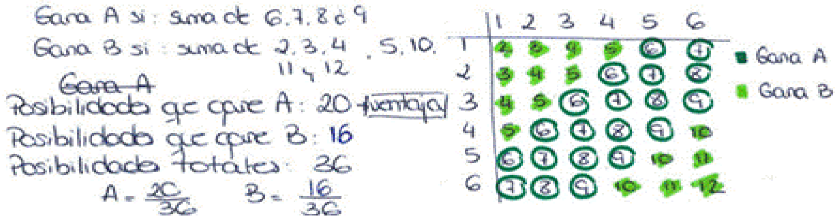


Figura 2. Respuesta correcta al apartado 1.

R2. Respuesta correcta relacional. Algunos estudiantes dan la respuesta correcta, calculando la probabilidad de ganar el juego los jugadores *A* y *B*, pero también indican que si se tiene en cuenta los resultados de la suma, sin considerar los diferentes resultados para obtenerla, aparentemente ganaría *B* y se cometería un error. Por tanto, ponen en relación los dos enfoques de la probabilidad. En el siguiente ejemplo (Figura 3), el estudiante analiza, en primer lugar, los resultados empíricos del juego, en donde *B* gana. Además, indica que este resultado lleva a cometer un error, pues en la probabilidad teórica, el que gana el juego es *A* con probabilidad 20/36 frente a *B*, con probabilidad 16/36. Esta última probabilidad está calculada correctamente.

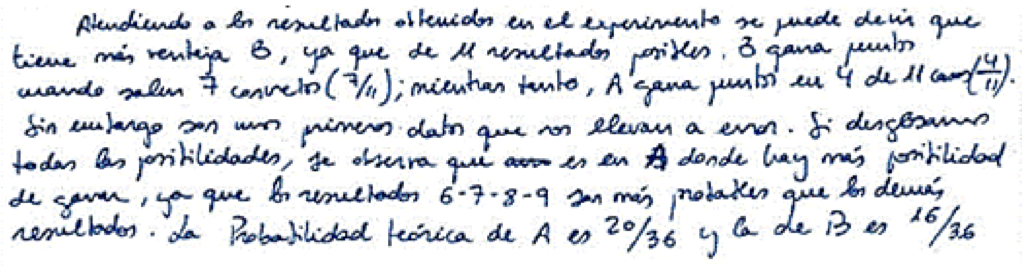

Atendiendo a los resultados obtenidos en el experimento se puede decir que tiene más ventaja B, ya que de 11 resultados posibles, B gana puntos cuando salen 7 concretos (7/11); mientras tanto, A gana puntos en 4 de 11 casos (4/11). Sin embargo son unos primeros datos que nos llevan a error. Si desglosamos todas las posibilidades, se observa que es en A donde hay más posibilidad de ganar, ya que los resultados 6-7-8-9 son más probables que los demás resultados. La Probabilidad teórica de A es 20/36 y la de B es 16/36

Figura 3. Respuesta correcta relacional al apartado 1.

R3. Parcialmente correcta. El estudiante afirma que tiene ventaja *A* sobre *B*, considerando que es más probable que al lanzar dos dados la suma de puntos sea favorable al jugador *A*. Sin embargo, el argumento utilizado, no es correcto. En el siguiente ejemplo (Figura 4), el estudiante da una respuesta correcta, pero su argumentación no lo es, ya que, por un lado, comete un error en el cálculo de probabilidades al considerar que los sucesos son equiprobables, es decir, presenta el sesgo de equiprobabilidad, (Lecoutre, 1992). Por ello deduce que *A* tiene una probabilidad de ocurrencia de 4/11, un 36% aproximadamente, y *B* de 7/11, un 64%. Añade que es difícil o raro que salgan dos números pequeños o dos grandes en el lanzamiento de los dados, sesgo de razonamiento que no hemos encontrado descrito en la literatura previa.

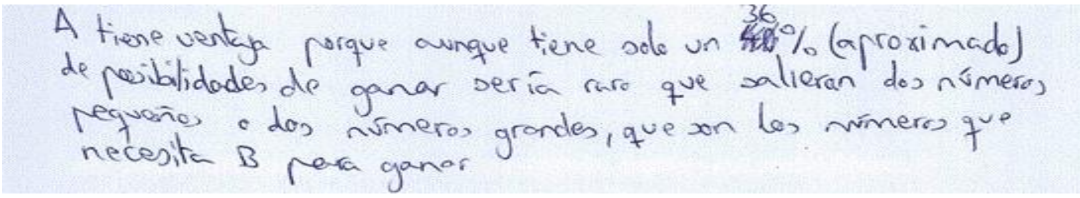
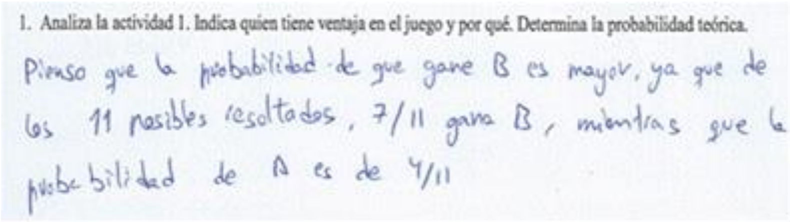

A tiene ventaja porque aunque tiene solo un ~~40~~ 36% (aproximado) de posibilidades de ganar sería raro que salieran dos números pequeños o dos números grandes, que son los números que necesita B para ganar

Figura 4. Respuesta parcialmente correcta al apartado 1.

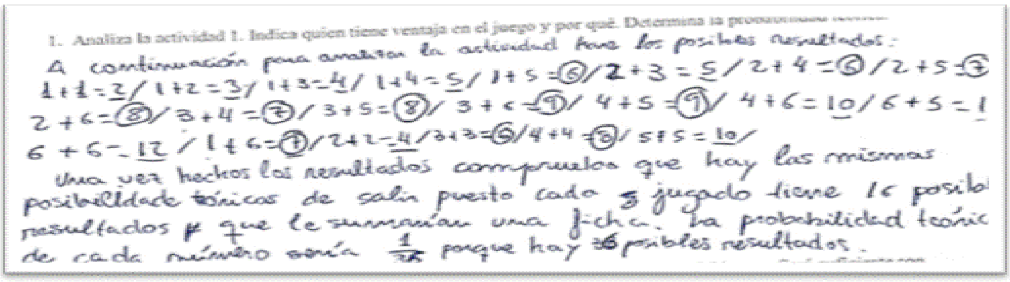
R4. Incorrecta por sesgo de equiprobabilidad. Algunos estudiantes indican que *B* tiene ventaja, pues consideran que todas las posibles sumas de los dos dados tienen igual probabilidad de ocurrencia, es decir, son consideradas sucesos equiprobables (sesgo de equiprobabilidad, Lecoutre, 1992). La diferencia con el caso anterior es que también fallan en identificar el jugador que tiene ventaja en el juego. En el siguiente ejemplo (Figura 5), el participante considera que las sumas de puntos al lanzar dos dados (del 2 al 12) son sucesos equiprobables. No forma el espacio muestral producto para analizar cuáles sucesos de este espacio conducen a cada posible suma. Es decir, considera todas las sumas con igual probabilidad de ocurrencia, lo que lleva a contestar erróneamente esta pregunta, considerando que gana *B* con probabilidad 7/11 frente a *A* que obtiene probabilidad 4/11.


1. Analiza la actividad 1. Indica quien tiene ventaja en el juego y por qué. Determina la probabilidad teórica.

Pienso que la probabilidad de que gane B es mayor, ya que de los 11 posibles resultados, 7/11 gana B, mientras que la probabilidad de A es de 4/11

Figura 5. Respuesta incorrecta por sesgo de equiprobabilidad al apartado 1.

R5. Incorrecta por espacio muestral reducido. Algunos estudiantes consideran dos sumas que aparecen en distinto orden como un único suceso. Este error, que también aparece con frecuencia en la resolución de problemas combinatorios, fue denominado por Navarro-Pelayo, Batanero y Godino (1997) como *error de orden*. En el caso que nos ocupa es comprensible, pues la suma tiene la propiedad asociativa; pero a pesar de ello, hay que diferenciar los resultados según el orden en los dados para el cálculo de probabilidades. Si se comete el error, el espacio muestral queda reducido, lo que lleva a responder de manera incorrecta a esta pregunta; como se muestra en el siguiente ejemplo (Figura 6).


1. Analiza la actividad 1. Indica quien tiene ventaja en el juego y por qué. Determina la probabilidad teórica.

A continuación para analizar la actividad haré los posibles resultados:
1+1=2 / 1+2=3 / 1+3=4 / 1+4=5 / 1+5=6 / 2+3=5 / 2+4=6 / 2+5=7
2+6=8 / 3+4=7 / 3+5=8 / 3+6=9 / 4+5=9 / 4+6=10 / 6+5=11
6+6=12 / 1+6=7 / 2+2=4 / 3+3=6 / 4+4=8 / 5+5=10 /
Una vez hechos los resultados compruebo que hay las mismas posibilidades teóricas de salir puesto cada jugador tiene 16 posibles resultados que le sumarían una ficha. La probabilidad teórica de cada número sería 1/36 porque hay 36 posibles resultados.

Figura 6. Respuesta incorrecta por espacio muestral reducido al apartado 1.

En la Tabla 2, se presentan las respuestas a este apartado de la tarea. Podemos observar que contestaron correctamente esta pregunta 43 estudiantes de los 60 (35% de los estudiantes en forma individual y el 90% de quienes responden en pareja), pero solo uno trata de relacionar explícitamente los dos significados de probabilidad. En las respuestas incorrectas, el error detectado con mayor frecuencia fue el sesgo de equiprobabilidad, presente en 12 estudiantes de la muestra (50% de los estudiantes que responden de modo individual y el 5% de quien responde en pareja), a los que hay que añadir tres estudiantes, que contestaron en forma parcialmente correcta, pero en su argumento cometieron el error de sesgo de equiprobabilidad. Solo un estudiante (5%) respondió incorrectamente a esta pregunta por error del espacio muestral reducido.

Tabla 2. *Frecuencia de respuestas de los estudiantes a la primera tarea*

Respuesta	Estudiantes	Parejas	Total de estudiantes
R1 Correcta	7	18	43
R2 Correcta relacional	1	0	1
R3 Parcialmente correcta	1	1	3
R4 Incorrecta por sesgo de equiprobabilidad	10	1	12
R5 Incorrecta por espacio muestral reducido	1	0	1

Al comparar los resultados de esta pregunta con la experiencia de aula de Rivas y Godino (2015), hacemos notar que en ambos trabajos predominan los mismos errores, que son el sesgo de equiprobabilidad y el espacio muestral reducido. Los autores no indican la frecuencia en sus estudiantes; en nuestro trabajo, contestaron correctamente a este apartado el 71,6% aproximadamente, el 20% de los participantes manifestó el sesgo de equiprobabilidad, y uno estudiante el error de espacio muestral reducido. Finalmente, tres estudiantes (5% de la muestra), contestaron de manera parcialmente correcta a esta pregunta.

Estimación del valor esperado mediante el enfoque frecuencial

2. Si realizáramos muchas tiradas, ¿Con cuánta frecuencia esperarías que saliera cada una de las sumas? ¿Con cuánta frecuencia ganaría *A* o *B*?

La frecuencia esperada de un suceso *S* en *n* repeticiones de un experimento viene dada por $Fe(S) = n \cdot P(S)$. En el apartado 1 se calcularon las probabilidades teóricas P(*A*)= 0,556 y P(*B*)= 0,444. Por tanto, en 1000 repeticiones, la frecuencia esperada sería que *A* ganaría 556 veces y *B* 444 veces, aproximadamente. Por otro lado, si se piensa en lanzar los dos dados 36 veces, y teniendo en cuenta los casos favorables para cada suma, se deduce que *A* ganaría 20 veces de 36 frente a 16 de cada 36 que ganaría *A*, o lo que es igual a 5 veces de 9 para *A* frente a 4 veces de 9 para *B*. Se encontraron las siguientes categorías de respuestas a esta actividad:

R1. Respuesta correcta. La mayoría de los estudiantes utilizaron los resultados de la pregunta 1 para responder a este apartado. Se da un valor correcto para la frecuencia esperada de resultados en que gana *A* y *B*, utilizando uno de los dos métodos expuestos anteriormente. Observamos este tipo de respuesta en el siguiente ejemplo (Figura 7), en que un estudiante primero presenta la probabilidad teórica con que gana *A* y *B* (deducida del cálculo de la probabilidad en el apartado 1). También afirma que la frecuencia esperada es 5 de cada 9 para que gane *A* y la de *B* es 4 de cada 9; por tanto, ha usado el cálculo de posibilidades del primer apartado.

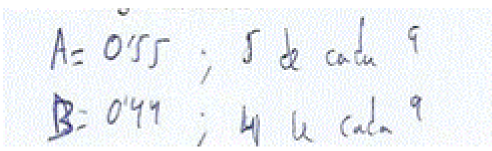


Figura 7. Respuesta correcta al apartado 2.

R2. Incorrecta por sesgo de equiprobabilidad. Como la mayoría de los estudiantes consideró la respuesta de la pregunta 1 para responder a este apartado, si cometieron el sesgo de equiprobabilidad en la primera pregunta, lo siguen manteniendo en este apartado, como podemos observar en el ejemplo (Figura 8) que se reproduce a continuación.

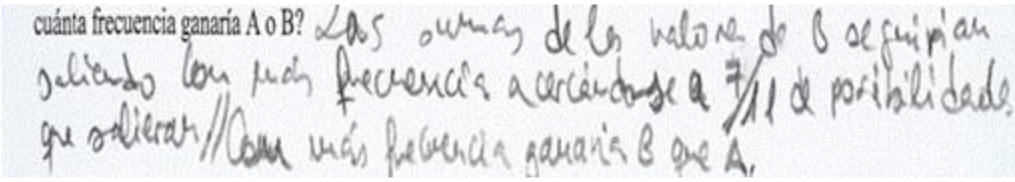


Figura 8. Respuesta incorrecta por sesgo de equiprobabilidad al apartado 2.

R3. Incorrecta por espacio muestral reducido. Algunos estudiantes que tuvieron este error en el primer apartado lo mantienen en este otro, lo que lleva a responder de manera incorrecta a esta pregunta. Un ejemplo es el siguiente, en el que el estudiante utiliza su respuesta de la pregunta 1, donde solo considera los pares (3,4) y (5,4) y no (4,3) y (4,5). Al omitir posibles sumas, su espacio muestral queda reducido a sólo 11 posibilidades (cuando, sin tener en cuenta el orden deberían ser 21 y 36 teniéndolo en cuenta). El estudiante muestra una falta de capacidad de enumeración sistemática, error típico del razonamiento combinatorio de acuerdo a Navarro-Pelayo, Batanero y Godino (1997).

En la Tabla 3 se presenta la frecuencia de estudiantes que respondió al apartado 2 de la tarea, en que contestaron correctamente 21 estudiantes de los 60, es decir, sólo la tercera parte y no hay respuestas parcialmente correctas, de lo que se deduce que esta pregunta fue difícil. El error más frecuente fue el sesgo de equiprobabilidad, presente en 10 estudiantes de la muestra y 4 estudiantes contestaron incorrectamente por espacio muestral reducido. Hubo muchos estudiantes (25 de 60) que no respondieron, confirmando la dificultad de la misma. Rivas y Godino (2015) no la plantean.

Tabla 3. *Respuestas de los estudiantes a la segunda tarea*

Respuesta	Estudiantes	Parejas	Total de estudiantes
R1 Correcta	7	7	21
R2 Incorrecta por sesgo de equiprobabilidad	10	0	10
R3 Incorrecta por espacio muestral reducido	2	1	4
R4 No responde	1	12	25

Comparación de los significados clásico y frecuencial de la probabilidad

3. Analiza la Actividad 2 ¿Piensas que la tabla con los datos de toda la clase es suficiente para sacar conclusiones del experimento realizado y las cuestiones planteadas? ¿Podríamos tomar como probabilidad de ganar *A* la frecuencia relativa de veces que gana *A*?

En esta pregunta se pretende que el estudiante reflexione sobre la diferencia que se ha observado entre el valor teórico de la probabilidad y la frecuencia relativa de los datos en 100 repeticiones del experimento. Se espera que el estudiante detecte la gran diferencia y rechace la posibilidad de estimar la probabilidad de ganar *A* mediante la frecuencia relativa. El estudiante debe responder que no serían suficientes los datos de esta tabla. La frecuencia relativa obtenida en el experimento es una estimación mala, en este caso, de la probabilidad. Las respuestas obtenidas fueron clasificadas de acuerdo a las siguientes categorías:

R1. Respuesta correcta. El participante responde correctamente a este apartado, al señalar que no es suficiente la tabla con los datos de toda la clase para obtener conclusiones del experimento realizado, porque observa que la frecuencia relativa obtenida en el experimento es una estimación errónea de la probabilidad. En el siguiente ejemplo (Figura 9), el estudiante indica que los resultados obtenidos en la tabla no son suficientes para poder sacar conclusiones del experimento realizado y expone que el número de veces que gana *B* en dicha tabla es superior al de *A*, con lo que observa que la probabilidad teórica y su estimación empírica no coinciden. Por lo tanto, no se puede considerar aceptable, pues da vencedor del juego a *B* y no a *A*, que sería lo correcto.

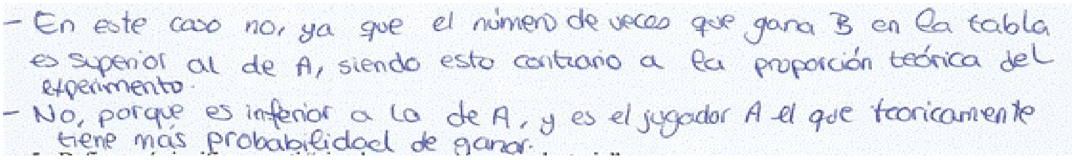

- En este caso no, ya que el número de veces que gana B en la tabla es superior al de A, siendo esto contrario a la proporción teórica del experimento.
- No, porque es inferior a la de A, y es el jugador A el que teoricamente tiene más probabilidad de ganar.

Figura 9. Respuesta correcta al apartado 3.

R2. Parcialmente correcta. Se considera en esta categoría de respuesta, aquella en que el estudiante responde correctamente una de las dos preguntas planteadas en el apartado pero no ambas. En el ejemplo que sigue (Figura 10), el estudiante considera que las tiradas de los dados en la tabla son suficientes para determinar un ganador en el juego. Pero, por otro lado, responde correctamente que no se debe considerar la probabilidad de ganar *A* según la frecuencia relativa, pues ésta indica solo las veces que gana *A* y no la probabilidad. Ello supone una contradicción entre las respuestas a las dos partes, de la que el estudiante no es consciente, lo que muestra su poca capacidad argumentativa.

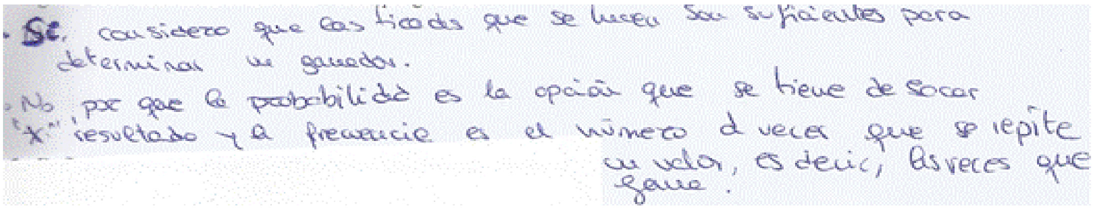

· Sí, considero que las tiradas que se hacen son suficientes para determinar un ganador.
· No, porque la probabilidad es la opción que se tiene de sacar "x" resultado y la frecuencia es el número de veces que se repite un valor, es decir, las veces que gana.

Figura 10. Respuesta parcialmente correcta al apartado 3.

R3. Respuesta incorrecta. Serían los estudiantes que responden incorrectamente las dos preguntas del apartado, como el siguiente ejemplo (Figura 11), donde el estudiante considera que las dos respuestas deben ser afirmativas. Por un lado, cree que es suficiente sacar conclusiones del experimento realizado con 100 tiradas. Por otro lado, considera que se puede tomar como probabilidad de ganar *A*, a la frecuencia relativa de veces que gana *A*, calculada a partir de los datos empíricos.

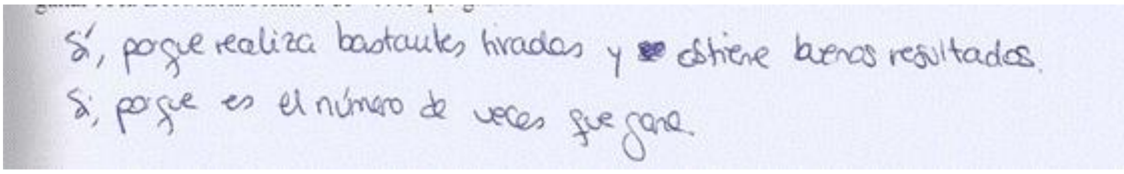


Figura 11. Respuesta incorrecta al apartado 3.

En la Tabla 4 se presenta la frecuencia de estudiantes que respondió al tercer apartado, de las cuales 48 fueron correctas. Esto muestra una buena comprensión de la relación entre la probabilidad en sentido clásico y su estimación frecuencial. Encontramos respuestas parcialmente correctas de seis estudiantes, que, aunque responden bien una de las preguntas, son inconsistentes en su argumentación. Hay también cinco respuestas incorrectas; finalmente, solo un estudiante no responde a esta pregunta, aludiendo que no sabe la respuesta. Los resultados son mejores en aquellos que trabajan en parejas.

Tabla 4. *Respuestas de los estudiantes a la tercera tarea*

Respuesta	Estudiantes	Parejas	Total de estudiantes
R1 Correcta	8	20	48
R2 Parcialmente correcta	6	0	6
R3 Incorrecta	5	0	5
R4 No responde	1	0	1

VI. Conclusiones

En nuestro estudio, resultaron sencillas la primera pregunta, en que los estudiantes determinan cuál de los dos jugadores tiene ventaja, calculando la probabilidad en el enfoque clásico, y la última en la que ponen en relación las definiciones clásicas y frecuencial de la probabilidad, pues la mayoría de los futuros profesores las resuelve correctamente. Fue mucho más difícil estimar la frecuencia esperada de veces que ocurre un suceso, en la segunda pregunta, lo que indica que estos futuros profesores no comprenden con profundidad la ley de los grandes números, ni siquiera en forma intuitiva, ya que sólo un futuro profesor ofrece un argumento relacionando ambos enfoques de probabilidad en la segunda parte de la actividad 1.

Al comparar los resultados con la experiencia en aula de Rivas y Godino (2015), podemos hacer notar que en ambos trabajos predominan los mismos errores en el cálculo de probabilidades, que son el sesgo de equiprobabilidad y el cálculo del espacio muestral reducido. También aparece la heurística de representatividad (Kahneman et al., 1982) cuando los estudiantes dan respuestas sin comprender la ley de los grandes números, sino que lo hacen en base a los resultados obtenidos en el juego empírico, a pesar de que a nivel declarativo (cuando se les pregunta) deciden que la muestra es de tamaño pequeño para obtener conclusiones. Se confirma, de este modo, la existencia en los futuros profesores de educación primaria de la heurística de la representatividad, encontrada en investigaciones previas (Azcárate, 1995; Serrano, 1996; y Batanero, Cañizares y Godino, 2005). Batanero, Cañizares y Godino (2005) indican que sus estudiantes superan en parte estos sesgos mediante actividades de simulación.

Aunque nuestros estudiantes habían llevado a cabo alguna de estas actividades, se realizaron en grupo (sin interacción personal con el ordenador, que sí hubo en el caso de Batanero et al., 2005). Deducimos que sería necesario un mayor tiempo dedicado a estas actividades y en lo posible, permitir la interacción de los estudiantes con la tecnología en la realización de las simulaciones.

Todos estos resultados nos plantean una problemática en la formación de los profesores, que se supone han de enseñar la probabilidad bajo estos dos enfoques en la educación primaria. Se necesitaría insistir más sobre este tema, para preparar mejor a los profesores para su futura labor docente.

Bibliografía

[1] Azcárate, P. El conocimiento profesional de los profesores sobre las nociones de aleatoriedad y probabilidad. Tesis doctoral. Universidad de Cádiz. 1995.
[2] Batanero, C. Significados de la probabilidad en la educación secundaria. Revista Latinoamericana de Investigación en Matemática Educativa (2005), 8(3), 247- 263.
[3] Batanero, C.; Cañizares, M. J.; Godino, J. Simulation as a tool to train preservice school teachers. En J. Addler (ed.), Proceedings of ICMI First African Regional Conference. [CD-ROM], Johannesburgo: International Commission on Mathematical Instruction, 2005.
[4] Batanero, C.; Henry, M.; Parzysz, B. The nature of chance and probability. En G. A. Jones (ed.), Exploring probability in school: challenges for teaching and learning, New York: Springer, 2005, p. 15–37.
[5] Bernoulli, J. Ars Conjectandi- 4ème partie (N. Meunier, Trans.) Rouen: IREM. (Original work published in 1713), 1987.
[6] Gal, I. Towards “probability literacy” for all citizens: Building blocks and instructional dilemmas. En G. A. Jones (ed.), Exploring probability in school: Challenges for teaching and learning, Nueva York: Springer, 2005, p. 39–63.
[7] Godino, J. D.; Batanero, C.; Cañizares, M. J. Azar y probabilidad. Fundamentos didácticos y propuestas curriculares. Síntesis. 1987.
[8] Gómez, E. Evaluación y desarrollo del conocimiento matemático para la enseñanza de la probabilidad en futuros profesores de educación primaria. Tesis doctoral. Universidad de Granada. 2014.
[9] Kahneman, D.; Slovic, P.; Tversky, A. (eds.) Judgment under uncertainty: Heuristic and biases. New York: Cambridge University Press, 1982.
[10] Konold, C. Informal conceptions of probability. Cognition and Instruction (1989) 6, 59-98.
[11] Laplace P. S. Théorie analytique des probabilités [Analytical theory of probabilities]. Paris: Jacques Gabay. (Original work published 1814), 1995.
[12] Lecoutre, M. P. Cognitive models and problem spaces in “purely random” situations. Educational Studies in Mathematics (1992) 23(6), 557-568.
[13] Ministerio de Educación, Cultura y Deporte (MECD). Real Decreto 126/2014, de 28 de febrero, por el que se establece el currículo básico de la Educación. Autor. 2014.
[14] Mises, R. von. Probabilidad, estadística y verdad. Madrid, España: Espasa-Calpe [Trabajo original publicado en 1928], 1952.
[15] Mohamed, N. Evaluación del conocimiento de los futuros profesores de educación primaria sobre probabilidad. Tesis doctoral. Universidad de Granada. 2012.